

ILLUSTRATIVE APPLICATIONS OF AIR TRAFFIC CONTROL SYSTEM CAPACITY STUDY METHODOLOGY

G. RAISBECK, J. L. EVERETT, and B. O. KOOPMAN

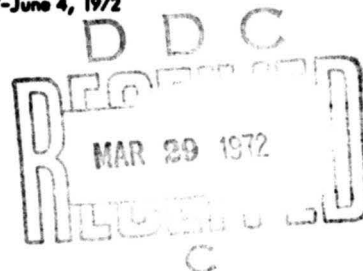
ARTHUR D. LITTLE, INC.
Cambridge, Massachusetts



U.S. International Transportation Exposition
Dulles International Airport
Washington, D.C.
May 27-June 4, 1972

NOVEMBER 1971

Interim Report



Availability is unlimited. Document may be released
to the National Technical Information Service,
Springfield, Virginia, 22151, for sale to the public.

Prepared for

**DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
SYSTEMS RESEARCH AND DEVELOPMENT SERVICE
Washington, D.C., 20591**

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**
Springfield, Va 22151

46

R

**Best
Available
Copy**

100

ACCESSION for

CPSTI	WHITE SECTION	<input checked="" type="checkbox"/>
99C	BAFF SECTION	<input type="checkbox"/>
4444444444		<input type="checkbox"/>

4444444444

1. Report No. FAA-RD-71-113		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Illustrative Applications of Air Traffic Control System Capacity Study Methodology				5. Report Date November 1971	
				6. Performing Organization Code 71860	
7. Author(s) G. Raisbeck, J.L. Everett, B.O. Koopman				8. Performing Organization Report No.	
9. Performing Organization Name and Address Arthur D. Little, Inc. Cambridge, Mass. 02140				10. Work Unit No.	
				11. Contract or Grant No. FA-70WA-2141	
				13. Type of Report and Period Covered Interim Report (Nov. 1970-Sept. 1971)	
12. Sponsoring Agency Name and Address Systems Research and Development Service Federal Aviation Administration Dept. of Transportation, Washington, D.C.				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract The long-range objective of this program is to develop tools and techniques to define, measure, and predict the capacity of an air traffic control system, which can then be used in analytical studies in support of long-range plans, management decisions, and system performance evaluations. The method of approach in this contract provides for testing and refining these tools by using them in typical current problems. This report illustrates the application of these tools to five typical current problems: (1) can we distinguish delays which are a normal consequence of the air traffic control system structure and the environment in which it operates from those delays attributable to some failure or deficiency? (2) in how much detail must the server statistics be described in a queuing model of terminal area delay? (3) what criteria for the collection of operational measurements are inferred from analytical methods? (4) how can a large number of factors affecting air traffic control system capacity be classified, for purposes of structuring cost benefit analyses, into a small number of classes, each class containing those measures which would be evaluated from a common point of view? and (5) how can quantitative safety goals for air traffic control subsystems and parts of aviation operations be derived rationally and practically from an overall system safety goal, and in what terms should these goals be expressed?					
17. Key Words Air Traffic Control Safety Capacity Non-stationary Queue Queuing System Modelling Demand Delay				18. Distribution Statement Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia, 22151, for sale to the public.	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 52	
				22. Price \$3.00	

TABLE OF CONTENTS

	Page
1. INTRODUCTION	1
1.1 Background, Purpose, and Scope	1
1.2 Summary	2
2. THE ORIGINS OF DELAYS IN AVIATION OPERATIONS	5
2.1 Statement of Problem and Summary of Conclusions	5
2.2 The Nature of Delay	6
2.3 Safety	7
2.4 The Demand	9
2.5 The Servicing System	9
2.6 Environmental Constraints	10
2.7 Friction-Producing Interactions Between Demand and Control	11
2.8 The Benefits of Delay	12
2.9 Reformulation of the Delay Issue	12
2.10 Measurement of Delay as a Tool for Diagnosis and System Improvement	12
3. INSENSITIVITY OF QUEUE PARAMETERS TO SERVER STATISTICS	15
3.1 Statement of Problem and Summary of Conclusions	15
3.2 Method of Approach	16
3.3 Server Statistics	19
3.4 Numerical Results	20
3.5 Conclusions	20
4. OPERATIONAL MEASUREMENTS IN THE TERMINAL CONTROL AREA	21
4.1 Statement of Problem	21
4.2 General Problems in Operational Testing	22
4.3 Quantities to be Measured Precisely	23
4.4 Qualitative Data to be Recorded	23
4.5 Deductions	24

TABLE OF CONTENTS (Continued)

	Page
5. STRUCTURING COST-BENEFIT METHODOLOGY FOR ATC CAPACITY IMPROVEMENT MEASURES	25
5.1 Statement of the Problem and Description of Results	25
5.2 Historical Background	25
5.3 System Improvements and Their Effectiveness	27
5.4 Results	31
6. ESTABLISHING PRACTICAL SAFETY GOALS IN AIR TRAFFIC CONTROL	39
6.1 Statement of the Problem	39
6.2 Method of Approach	40
6.3 Conclusions	41
REFERENCES	43

LIST OF FIGURES

Figure No.		Page
2.1	Hourly Arrival Rates — Terminals A and B	8
3.1	Queue Statistics under Various Server Assumptions — Expected Number in Queue	17
3.2	Queue Statistics under Various Server Assumptions — Standard Deviation of Number in Queue, Probability of Empty Queue, and Probability of Full Queue	18
5.1	Air Traffic Control Improvement Pattern	29
5.2	Limiting Airport Characteristics Related to Key Effects	32
5.3	Limiting Aircraft Characteristics Related to Key Effects	33
5.4	Limiting ATC Characteristics Related to Key Effects	34

LIST OF TABLES

Table No.		Page
2.1	Sources of Friction in the Aviation System by Cause	11
5.1	Capacity-Influencing Items Considered as Equivalent to Other Items	35
5.2	Capacity-Influencing Items Omitted for Current Structural Irrelevancy	35
5.3	Relationships between ATC Improvement Measures and Their Key Effects upon the System	36

1. INTRODUCTION

1.1 BACKGROUND, PURPOSE, AND SCOPE

This report is one of a series describing work on a continuing study of air traffic control (ATC) system capacity. The FAA frequently undertakes analytical and experimental studies in support of long-range plans, management decisions, and system performance evaluation. The objective of this program is to develop methods to define, measure, and predict the capacity of an ATC system, which can be used in such studies. An essential part of this work is to give a precise meaning to the descriptive term "air traffic control system capacity." Another part is to find quantitative relationships between the capacity of the air traffic control system and performance of the air transportation system of which it is a part. A further important part is to find quantitative relationships between the capacity of the ATC system and the characteristics of the elements -- equipment, procedures, people, and configuration -- of which it is comprised. To carry out this program, we are defining a number of units, terms, and measures, developing methods for measurement, analysis, and computation, and using a variety of models of various subsystems.

The major product of this work, is therefore, technical tools to be used in further analytical studies by the FAA. The validity and usefulness of tools under development is most easily judged by practical application. Therefore, our method of approach in this program is to test and refine these tools by applying them to typical current problems. Five such illustrative applications are described in the following chapters.

We expect a number of benefits can be derived from conducting and publishing these brief studies. In the first place, the analysts working on this program have an opportunity to see for themselves what can be done with the terms, measures, definitions, and analytical and empirical measures being developed during the course of this work. Second, those who might ultimately benefit by applying these tools have an opportunity to see them in use in their formative stage, and in the concrete context of particular applications, to offer criticisms which might be difficult to anticipate and state in abstract terms. Third, it provides the FAA with an opportunity to react by directing our attention to classes of problems more important to them than those chosen for illustration.

None of these examples has been carried out on a scale broad enough to produce definite answers. The purpose of the work described in this report is to show how these tools can be used, to invite constructive criticism, and to arouse interest in further applications.

1.2 SUMMARY

Chapter 2 is concerned with the problem of distinguishing between two classes of ATC system delays: (1) those which are a normal consequence of the system design and the environment in which it operates, and (2) those attributable to failure or deficiency in the ATC system. Delays in aviation operations are investigated from the perspective of congestion theory. We discovered that some amount of delay is inevitable in any system where a large number of users with relatively loose coordination are creating demands for service from a relatively small number of servers. Like friction in a machine or resistance in an electrical circuit, this delay may be reduced but we cannot expect to abolish it entirely. A tabulation of commonly recognized sources of delay shows that most are, in this sense, normal and that only a very few can be described as failures or deficiencies. Thus we concluded that the question "at what level is further investment in delay reduction no longer repaid by the value of the decreased delay" was usually appropriate, whereas the questions "how can this delay be eliminated" and "who is responsible for this delay" were usually inappropriate. We also concluded that delay measurement might be a sensitive measure of ATC system loading, but it would usually be a poor tool for diagnosing the cause of delay.

Chapter 3 presents an investigation of the problem of how to make an analytical representation of server statistics suitable for analyzing terminal-area queues. When the amount and distribution of delay times become an issue, a mathematical queue is the appropriate idealization of air terminal operations for purposes of analysis and the interpretation of data. An important feature of any such representation is to include in it features corresponding to those aspects of reality which influence the answer, and to remove, for purposes of simplification, those features which have negligible effect on the result. In many classical stationary queuing problems, the expected value and fluctuations in queue delay are very sensitive to the statistical distribution of service times. In a number of examples representative of terminal area queues with large diurnal demand fluctuations, we found queue behavior to be quite insensitive to server statistics. When the demand rate fluctuates rapidly between values substantially above and well below the average service rate, the resulting dynamic changes in the queue are much greater than effects associated with server statistics. This observation is directly applicable to the analysis of terminal-area queues, because most practical terminal-area queuing problems involve large diurnal demand fluctuations, relatively dead periods at night, and rush-hour peaks in which demand temporarily exceeds average service rates. As a practical matter, it is not necessary either to measure or to make a faithful model of the server statistics. Any mathematically and computationally convenient representation may be adopted without strongly affecting the results.

Chapter 4 is concerned with the problem of how to make operational measurements in the terminal control area. It is impossible to make controlled measurements of operations with all meaningful combinations of system variables. It is hopelessly expensive and complex to record "all" measurable aspects of a real operation. Data should be collected in the context of a conceptual representation or model of the system under study, a hypothesis to be tested, and an experimental design for the organization and reduction of data. This methodology is illustrated with a very brief analysis of what should be measured at an airport operating under IFR conditions, why, and what information could be deduced from the results concerning the relation between distributions of inter-aircraft separations and terminal area capacity.

Chapter 5 is concerned with the problem of structuring cost benefit analyses involving a large number of factors, described in a companion report,¹ which affect ATC system capacity. A universal model which includes all of the factors will probably be complex and unwieldy. An analysis of such factors one at a time requires too many repeated attacks on the problem and fails to consider significant interactions. Without carrying out any specific cost benefit analysis, it is possible to chart an intermediate course. Some factors are combined because they represent only slightly different perspectives or processes which are fundamentally similar. Another group can be dropped from consideration as not being structurally relevant to the capacity question, although they did deal with subjects influencing ATC costs or a non-ATC system-related mechanism for the relief of capacity problems. The residual items were then grouped in terms of the various capacity bottlenecks they influenced, either singly or simultaneously. When laid out in a diagrammatic scheme, these were discovered to fall into a relatively small number of groups with few overlaps. For the purposes of cost benefit analysis, each one of these groups can be evaluated from a common point of view. Thus, the number of different approaches required is far less than would be suggested by the number of potential measures to be considered.

Chapter 6 presents a brief review of a more extensive companion report² on establishing practical safety goals in air traffic control. Because of an inverse relation between traffic densities and safety, some increase in traffic-handling capacity is always possible by lowering the safety margins. On the other hand, there is no way to take all of the risk out of their transportation. Therefore, some finite level of risk is implicit in any system capacity determination. In this work, we show how to select the unit for risk in terms of which an overall goal for air transportation safety can be stated. We then show a rational process whereby the total risk can be subdivided into fractions associated with various phases of flight and causes. The resulting partial risks are stated in quantitative terms suitable as design goals for planners of subsystems and operations. The use of a particular unit for the measurement of risk is demonstrated, and a particular methodology for establishing practical safety goals is illustrated. Several further conclusions are stated at the end of the chapter.

2. THE ORIGINS OF DELAYS IN AVIATION OPERATIONS

2.1 STATEMENT OF PROBLEM AND SUMMARY OF CONCLUSIONS

In everyday life, delay is frequently the consequence of error or human frailty. It is easy to jump to the conclusions that a delay in aviation operations is always the result of something going wrong, and that when defects are removed from the system, such delays will disappear. Such conclusions are unwarranted, and to assume their validity is to invite errors from incorrect inferences and unsound formulations of real problems. The purpose of this chapter is to forestall some of these errors by explaining the nature of delay from the point of view of queuing theory. It is not aimed at the solution of any specific problems, but at the avoidance of a whole class of problems.

By any standards, the national aviation system of the United States is a very effective means of transportation. Although its growth has been phenomenal, it has compiled an outstanding record in safety and in its abilities to meet the demand. With its benefits are associated some penalties and costs, and everyone with an interest in the system should recognize how these penalties and costs arise and should cooperate to ameliorate them as much as possible.

Delay is only one of the many costs and benefits which measure the effectiveness of the ATC system. Many sources, including the characteristics of the demand, interaction with the environment, economic and administrative constraints, and characteristics of the ATC system, can result in delay.

Queuing theory teaches that finite delay should be expected in the air transportation system, just as finite friction is expected in a mechanical machine or finite resistivity in an electrical circuit. Delay can be reduced at a price, but a condition of infinitesimal delay is not likely to give the best balance among costs and benefits in air transportation. When major sources of delay are tabulated, it is easy to see that most delays are not caused by air traffic control, even though air traffic control is often instrumental in providing the balance between delay and other penalties and in communicating the control decisions. Delay measurement may be a sensitive measure of system loading, but it will be a poor tool for diagnosing the cause of delay.

2.2 THE NATURE OF DELAY

Delays occur in any phase of life when demands for service arise from a multiplicity of sources acting more or less independently of one another. Life is full of examples. Automobiles are subject to traffic delays. Supermarket customers experience delays at the checkout counter. Patients spend more time than they wish in the doctor's waiting room, and sometimes many seconds go by before we hear the dial tone on the telephone.

Delays are annoying and can be costly. They are annoying for purely psychological reasons with which everyone is familiar. They often are associated with increased costs which may, in some cases, be very obvious and, in other cases, less so. If an unexpected delay causes a missed appointment and, therefore, the collapse of a business deal, the cost of the delay is both very evident and very painful. On the other hand, the extra 15 minutes required to cope with the rush hour may not be so painful and evident, although it could entail using extra gasoline and additional wear and tear on the automobile.

Since delays are so invariably disliked by everyone, except perhaps by condemned prisoners, the pressure exerted to reduce them is intense and continuous. The highway system responds by building more freeways. Supermarkets add more checkout booths, and the medical schools increase their enrollments. However, it is a common observation that delays are totally eliminated only very seldom.

The reasons why a delayless paradise can never be attained are rooted in the basic nature of the congestion phenomenon and its associated economics. These reasons dictate that reduction in delays be brought about at a price. The price to the user of the system might be encroaching upon his freedom to demand service at any time or place of his choosing, reduced quality of service, or the same service at higher price. On the other hand, the servicing system can respond by investment in more facilities, increase in efficiency, and the like. In any given case, a balance will almost never occur at zero delay, because zero delay implies total regulation of the demand (which is totally abhorrent to human nature), or at infinite cost to the service system, which is unfeasible. In some systems, as for example, the ATC system, zero delay would require total control of the environment which is, of course, impossible.

We all are aware that a totally frictionless machine can never be practically realized. However, we can approach it at the cost of resources in the form of time, effort, and money. How, in a particular case, we allocate these resources depends upon the sources of friction in the machine we are trying to improve. Similarly, a totally frictionless ATC system can only be approached. To see how best to allocate the resources available for this task, we must first identify and understand the sources of friction which prevent it from ever attaining absolute perfection. The objective of this report is to examine the anatomy of the congestion process and explain how three principal sources of friction – demand, the servicing system, and nature – contribute to the total friction.

2.3 SAFETY

In any discussion of the operations of an aviation system, the vital importance of safety must necessarily be recognized. Any compromises which present themselves in the operation of the system are always resolved in favor of safety.

Consequently, if the choice is between either increased throughput time or a decrease in the safety margin, the designer of the ATC system almost inevitably chooses the former. This choice is demonstrated by the fact that travel by air carrier within the United States is one of the safest means of travel available to the public.

2.4 THE DEMAND

The demand for service on any given module of the air traffic control system is characterized by:

- Wide fluctuations in magnitude,
- Variety in size and composition,
- Unpredictability, and
- Multiplicity of independent sources.

Figure 2.1 shows the demand pattern for service typical of the ATC system. The curves represent the sum of hourly arrival and departure operations at two major hub airports in the United States on one specific day. The demand is usually very small during the nighttime hours, but rises sharply to peaks which occur in the late afternoon. In the case of terminal A the ratio of peak to minimum is 10 to 1, and for terminal B, approximately 65 to 1. This type of demand is observable not only at airports, but in general throughout the entire ATC system. Its causes are multiple, but to a great extent are based on times people prefer to fly and times they do not.

The demand is variable in character. If, for example, we were to observe the situation at terminal A or B on any other day, the curves would be different. This difference might be large or small. Certainly, the sequence of aircraft types would be different. The total volume and the peaking would be different if the two days were, for example, a Friday and a Sunday. Moreover, the ratio of general aviation to air carrier types would not be the same.

Finally, the demand is unpredictable, except perhaps in terms of gross averages. Emergencies, potential conflicts, weather, equipment failures are, by their nature, highly unpredictable (except as averages), but they can often have drastic effects upon the volume and schedules of air traffic at any point in the ATC system.

One of the most important reasons why delays are variable and unpredictable has been mentioned already, namely, that the demand arises from a multiplicity of largely independent sources. A pilot may have a general idea of the traffic conditions he can expect on a given flight. However, the timing of his request for taxi instructions will be based much more on when his aircraft is ready

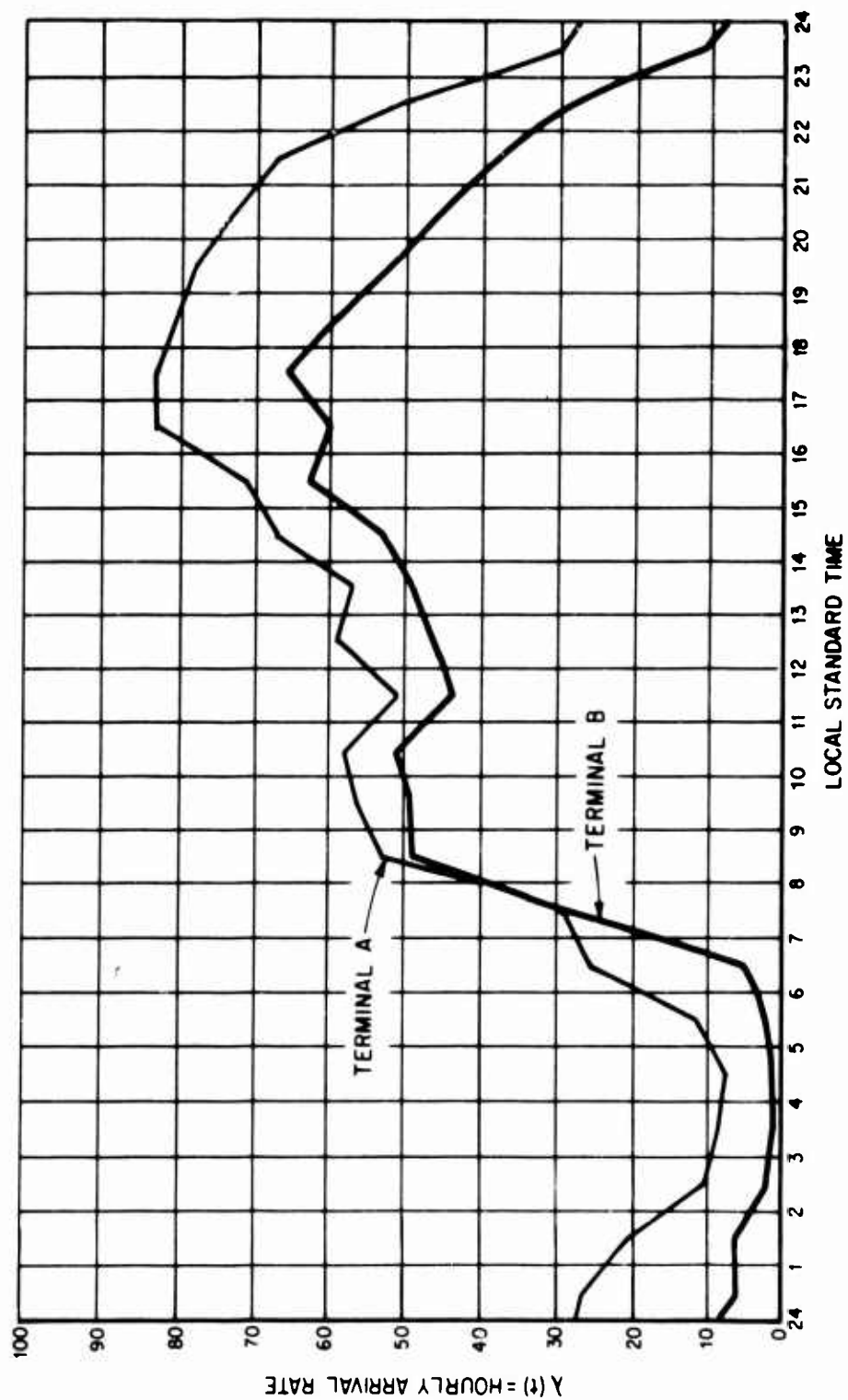


FIGURE 2.1 HOURLY ARRIVAL RATES - TERMINALS A AND B

for departure than upon the status of all the other aircraft which might be in those areas in which he may find himself during his forthcoming flight. Therefore, so far as the ATC system is concerned, his demand for service is largely independent of the demands made by other aircraft, and this applies to all the subsequent demands he will make throughout his journey. Consequently, the behavior of the ATC system will be describable not as a deterministic process telling the exact position and timing of each aircraft, but rather as a statistical phenomenon obeying probabilistic laws. Consequently, we have to deal with averages and distributions, and exact predictions are inherently impossible.

2.5 THE SERVICING SYSTEM

The ATC system may be thought of as a command and control system made up of links of sensors, navigation aids, man-machine teams, and the associated data processing, display, and communications. There is no point here in discussing its functions, except to say that it is subject to a demand for its services which has the characteristic discussed in the previous section. Because of the basic characteristics of this demand, namely, it is variable and non-predictable, there will be times at which the capacity of any given link will be stressed since, because of the costs involved, no link can conceivably handle all the possible demands to which it might be subjected.

On the other hand, even if the nominal capacity of a given link is large enough to handle the demand, nevertheless, the actual capacity may be less. Circumstances which could lead to such situations include equipment malfunction or degradation of controller team efficiency through lack of training or inadequate staffing. A third possible source of capacity attenuation might occur if, for some reason or other, the standard procedures were less than optimum for the particular conditions which existed in the particular link involved. Moreover, for various reasons which will be discussed in the next section, the ATC system is not always able to exploit all the capacity that is available at any given instant of time.

2.6 ENVIRONMENTAL CONSTRAINTS

One objective of the ATC system is to direct the movement of aircraft from one point to another in the most efficient manner, namely, as quickly and safely as possible. However, this operation takes place within an environment which, at any given time, imposes constraints on the actions of the ATC in pursuit of this objective. Those constraints come partly from the physical environment, namely, airports, weather, and the geography, while other constraints are administrative in nature, designed to enhance safety, or to reduce the undesirable impacts of air operations on other human activities. Examples of the latter are ATC regulations in the form of standard procedures (e.g., SID's and STAR's), noise abatement

rules, airway configurations, altitude and area restrictions, and the like. The net effect of all those constraints is to limit the scope of the actions which are available to the air traffic controllers in any given situation and, therefore, effectively to limit the throughput of aircraft in both the enroute and the terminal areas.

In absolute terms, the throughput limitations in a particular link are not necessarily the same from one time to another. One prime cause of this anomaly is the weather, the limiting effects of which on all aspects of air operations are too obvious to mention. Since weather is variable, the limitations it imposes are also variable. Moreover, weather, coupled with other environmental constraints such as runway and taxiway configurations, can impose drastic and highly variable constraints upon the effective throughput limitations of a particular terminal area at any given time.

2.7 FRICTION-PRODUCING INTERACTIONS BETWEEN DEMAND AND CONTROL

In the previous sections we discussed the nature of the demands imposed on the ATC system and the constraints under which this system operates. In this section, we will consider the interactions that take place between the demand and the control system. Our objective is to identify the sources of friction which inhibit the total system from performing in the most preferred manner, namely, without any delays in the operation.

The nature of the demand itself will cause friction. As we have seen, the major characteristics of this demand are that it is variable and unpredictable. Because no system can provide for all the possibilities that are statistically possible, inevitably there will be instances when the existing capacity will be strained and delays will occur. One of the principal characteristics of the demand that cause delays are the peaks which are considerably in excess of the average demand itself. We may be able to deal with these peaks most of the time, but the nature of the aviation system is such that from time to time a peak will occur which cannot be taken care of. We have also seen before, peaking and the demand will both cause delays. On the other hand, capability may be degraded through equipment malfunctions, degradation of controller performance, and errors in judgment on the part of the controllers. All these things lead again to delays to aircraft.

Finally, we have identified the many constraints which limit the ways that the ATC system can operate. Complying with these constraints often leads to a slowdown of the operations. Some of these constraints are typical in nature and come from the environment. Some of them are administrative and arise from such reasons as safety or noise abatement, and are essentially exogenous to the control system at any point in time.

Table 2.1 summarizes the various sources of friction in the system which lead to delays. They are categorized according to their source, namely, demand, ATC system, and constraints.

TABLE 2.1
SOURCES OF FRICTION IN THE AVIATION SYSTEM BY CAUSE

Demand	Air Traffic Control	Environmental and Administrative Constraints
Peak periods—variability	Equipment malfunction	Physical Weather
Unpredictability	Controller team proficiency	Runway configuration
Multiple independent sources	Non-optimum procedures	System Geographical configuration Sensors Communications Navigation aids Landing aids
Mix of users Aircraft types Military, carrier, general		
Emergency occurrence Conflicts Malfunctions		Administrative Authorized procedures STAR's SID's Noise abatement Safety
Variable pilot proficiency		
Variable services requested Preferred routes and altitudes		

2.8 THE BENEFITS OF DELAY

By allowing delay to buildup here and there, we can enable the system to operate under highly variable and largely unpredictable demand and environmental conditions. Delay is a relatively benign penalty, compared to cancellation or risk of collision, and can be spread around among many users in small amounts. A tremendous number of emergencies and hazards can be alleviated by skillful use of delay. Therefore, the option of subjecting some users to delay is a positive benefit to the national air transportation system.

Unfortunately, when any particular user is delayed, he usually does not know why, and he may wonder. Did weather restrict the use of certain airways? Did noise abatement procedures limit the landing rate at a terminal? Have controllers been diverted to assist an aircraft in genuine trouble? The user often

finds out about the delay from ATC in the form of denied or delayed clearance, a course diversion, a slowdown, or assignment to a holding pattern. As the bearer of bad news, either the controllers or the ATC system is blamed, no matter which of the sources listed in Table 2.1 is the ultimate cause.

2.9 REFORMULATION OF THE DELAY ISSUE

We have seen that the question "How can delay be eliminated?" is based on an erroneous idea of the role of delay. Moreover, the question "Who is responsible for my delay?" may very well yield an answer, but one which is unlikely to reduce the delay.

It is better to ask: Is the overall quality of service the best that can be achieved for the total system cost? Would an alternate configuration or set of operational procedures, which might result in more or less delay, produce a better standard of service? Is it worthwhile to make further investment to reduce delays and other penalties?

It is worth noticing that some aspects of these questions are being asked and answered. Flow control, for instance, reduces the unpredictability of demand at a terminal, but at the expense of denying service to some users. One can no longer find* a scheduled non-stop flight from Washington National Airport to Boston departing between 4:45 p.m. and 6:30 p.m., although there are four from 2:55 p.m. to 4:45 p.m. and four from 6:30 p.m. to 8:15 p.m. A balance has been struck between schedule preference and smoothing out the demand curve.

2.10 MEASUREMENT OF DELAY AS A TOOL FOR DIAGNOSIS AND SYSTEM IMPROVEMENT

Data are not organized on their own. Any actual set of data is a small subset of all the possible observations that could be made. It is advantageous to collect and process the least amount of data which can support the desired inferences. We must know how the data will be processed and what kind of inferences will be drawn from it before we can have any assurance that allows us to decide what data to select.

Therefore, the data requirements for research, development, and investment decisions are determined by the *analytical models* from which effectiveness is estimated. These analytical models are based upon a detailed understanding of the system and the interaction of its components which consist of both men and machines. Some of the most vital data needed can probably best be obtained by special tests suggested by the analyst rather than by an omni-purpose approach.

* During November 1971. Schedules are changed frequently.

Such an approach may, however, have value in revealing certain unsuspected factors affecting delay. Suspected factors, such as weather, runway conditions, demand patterns, and the like, are already understood, but additional statistics regarding them will be of considerable importance. However, data about delays alone are insufficient. As we have seen, delay is one of numerous performance parameters that are traded to find a condition of most satisfactory overall system performance. Data about the other performance parameters and about demand and environmental and administrative conditions which affect delays must be gathered also. If these data are gathered indiscriminately, the amounts will be excessive. It is better to pose a problem or a question first, next figure out a logical train of analysis which can answer the question or solve the problems, and only then decide what data to gather to support the analysis.

The ATC system operates as a multiply compensated, highly buffered, feedback system in which delay is one of the least costly penalties for system overload. The controllers will take any of the sources of friction (Table 2.1) and turn it into a delay as a way of avoiding a worse service penalty. Therefore, delay will be a sensitive index of the degree of loading of the system, but probably a very poor diagnostic tool for discovering the origins of the loading.

3. INSENSITIVITY OF QUEUE PARAMETERS TO SERVER STATISTICS

3.1 STATEMENT OF PROBLEM AND SUMMARY OF CONCLUSIONS

In many practical steady-state queuing problems, choice of a model for service-time distribution is an important issue. This is because the practically interesting details of queue behavior occur under near-saturation equilibrium conditions, when behavior is extremely sensitive to server statistics.

We have therefore investigated the problem of how to make a mathematical representation of server statistics suitable for analysis of terminal-area queues. These queues are not stationary, that is, the demands and service rates are functions of time rather than being constant. During some exploratory analysis with a number of models, some of them very detailed and correspondingly complex, interestingly enough, we discovered that the behavior of these queues was unlike the behavior of stationary queues in one important respect: they are quite insensitive to server statistics. The underlying reason is that terminal-area queues are driven by a widely fluctuating demand function. The demand rate does not dwell at a value close to the service rate long enough for near equilibrium to be reached and, therefore, the asymptotic queue behavior under these conditions is rarely significant. When the demand rate fluctuates rapidly between values substantially above and well below the average service rate, the resulting dynamic changes in the queue are much greater than effects associated with server statistics. As a practical matter, a mathematically and computationally convenient representation of server statistics may be chosen with considerable freedom without much affecting the results.

3.2 METHOD OF APPROACH

This problem and its solution are taken up in a previous report³ and are reviewed briefly here. We calculated a number of queue parameters, including expected number in the queue and its standard deviation and the probabilities of full or empty queues, for a single-server finite queue with a fluctuating demand and several assumptions about server statistics.

In particular, we used the same pair of demand curves used in previous reports.^{3,4} These were based on actual activity statistics at Kennedy and LaGuardia airports during one month of 1968. These activity data represent the most accurate and detailed sample of activity at very large airports which were readily available to us, and we have used them repeatedly as typical examples of diurnal demand variations. While rates of arrivals only are presumably about half as great as the total activity rates (on the average and at any particular time), the activity rates do give a realistic idea of proportional degrees of variability during the course of the 24-hour day of the actual arrival rates.

Preceding page blank

In the examples below, we have set the demand rate $\lambda(t)$ equal to the average daily activity rate at Terminal A (J.F. Kennedy Airport), as shown in Figure 2.1.

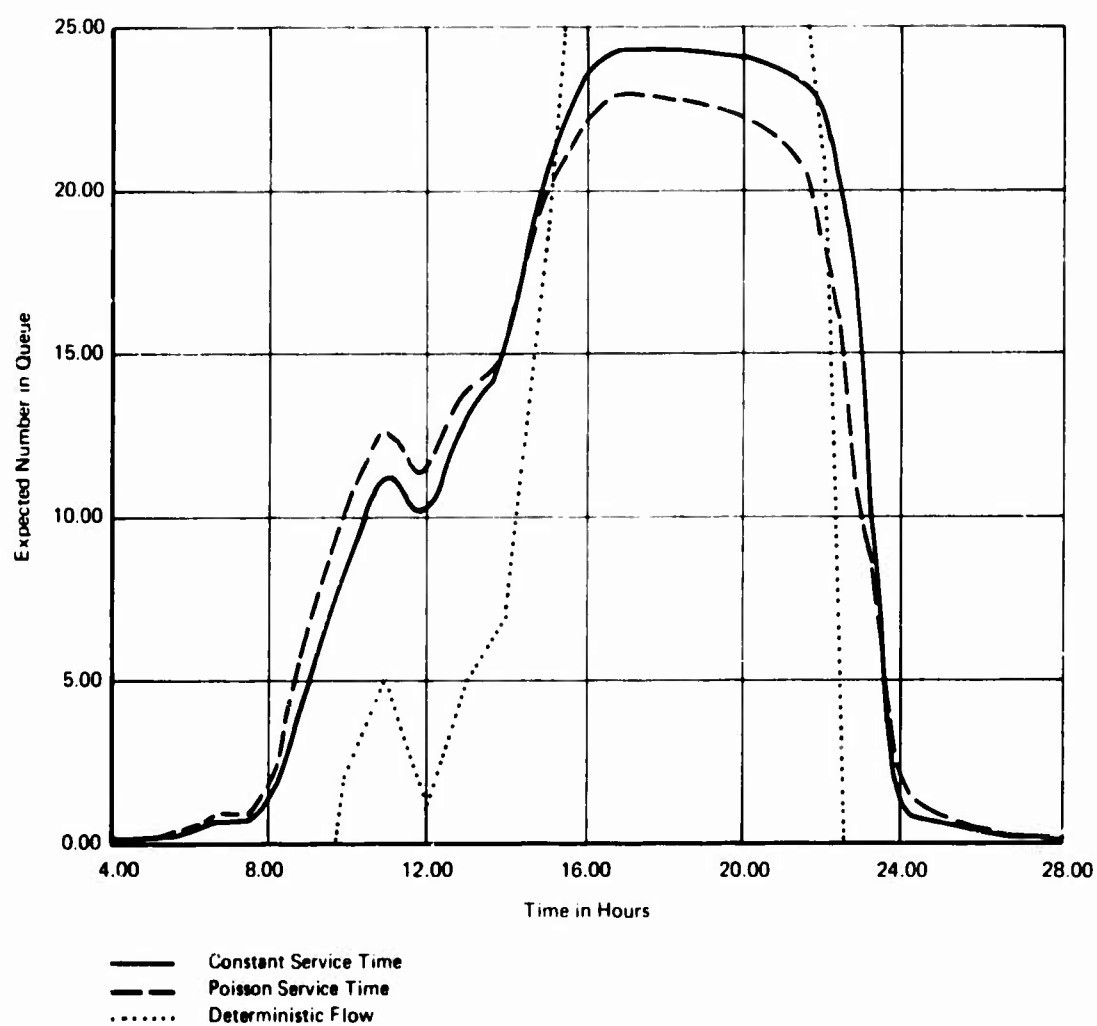
The maximum number allowed in the queue is assumed in each case to be $m = 25$. It is obviously somewhat unrealistic to combine the large arrival rates introduced in the last paragraph with this rather limited landing facility. For example, a simple count shows some 60 spaces for stacking available at J.F. Kennedy Airport. Nevertheless, the assumption is useful in bringing out most vividly the effects of congestion and the improvements obtainable by even moderate increases in service rates, as well as in flow control.

3.3 SERVER STATISTICS

In these examples, a service rate (μ) of 55 operations per hour was assumed. Three different sets of assumptions about demand and server statistics were made. The first assumption (solid line on Figures 3.1 and 3.2) is that the demand is a Poisson process with a time-varying rate defined by Figure 2.1, and that the service time is uniformly constant. In real life, this corresponds to the assumption that every landing or take-off occupies the runway for precisely the same time, and that no time is lost from delivery in accuracy or otherwise. Of all assumptions about service having a constant rate, this assumes the greatest possible regularity.

The second set of assumptions (dash lines on Figures 3.1 and 3.2) is the same with respect to the demand, and assumes that service times are generated by a Poisson process having a constant average of 55 operations per hour. On this assumption, interservice times have an exponential distribution with an average interval of 65.45 seconds, but with a service time of less than 1 second or greater than 5 minutes over 1% of the time. For a real airport operation, this distribution is absurd, because interarrival times as low as 1 second could not be tolerated for safety reasons and a runway would never be exclusively occupied for 5 minutes, except in case of an accident. However, this probability distribution defines the most random set of server statistics which is consistent with a stationary average. It was chosen not as a representation of airport operations, but to illustrate the fact that queue behavior is insensitive to this choice.

The third set of assumptions (dotted lines in Figures 3.1 and 3.2) is for a deterministic flow, where the server behavior is the same as that in the first example and the demand is fully deterministic; that is, both the demand and the service are treated as continuous fluid flow with deterministic rates, rather than as quantized flow with random fluctuation. This degree of regularity would be closely approximated in real life if arrivals were perfectly scheduled with an interarrival time equal to the reciprocal of the time-varying demand curve of Figure 2.1.



TERMINAL "A" $\mu = 55$ PER HOUR = $1/c$

FIGURE 3.1 QUEUE STATISTICS UNDER VARIOUS SERVER ASSUMPTIONS – EXPECTED NUMBER IN QUEUE

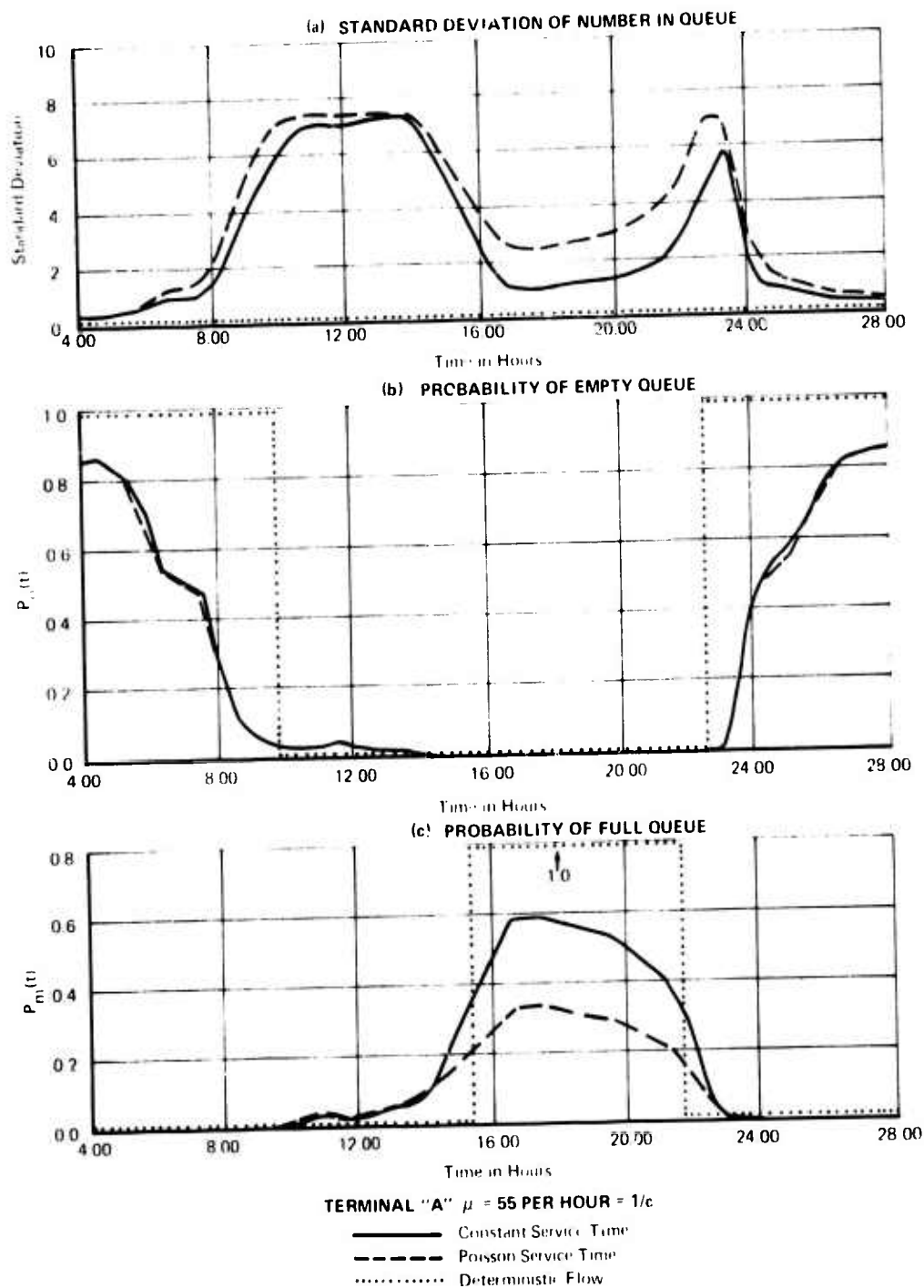


FIGURE 3.2 QUEUE STATISTICS UNDER VARIOUS SERVER ASSUMPTIONS—STANDARD DEVIATION OF NUMBER IN QUEUE, PROBABILITY OF EMPTY QUEUE, AND PROBABILITY OF FULL QUEUE

3.4 NUMERICAL RESULTS

Figure 3.1 shows the expected number in the queue as a function of time of the day for each of the three conditions. Figures 3.2a, b, and c show, respectively, the standard deviation of the expected number in the queue, the probability of an empty queue, and the probability of a full queue for the same three conditions.

It should be noted that the solid and the dashed curves in Figure 3.1 which represents the expected number in the queue under the assumptions of constant service time and Poisson service statistics, respectively, are fairly close for their whole length. Inasmuch as waiting time is also proportional to the expected number in the queue, we can conclude that the distribution and average of waiting times under these two service assumptions are also nearly equal.

Figure 3.2a shows the standard deviation of the estimate in Figure 3.1 and is a measure of the fluctuations to be expected around the mean value. In the case of the solid curve, the fluctuations are due to the randomness of the demand only, for the service is completely deterministic. On the other hand, in the case of the dashed line, both the demand and the service contribute to the fluctuation. As one would expect, the standard deviation of the expected number in the queue is always somewhat greater under the assumption of Poisson service statistics than it is under the assumption of constant service time. However, it is noteworthy that this difference is rather small. Also, over most of the day the standard deviation of the estimate of the expected number in queue (Figure 3.2a) is of the same order of magnitude as the difference between the estimates for constant and Poisson service times, respectively (Figure 3.1), or less. Therefore, we can say that the intrinsic uncertainty in the estimate in Figure 3.1 is as great as, or greater than, the difference between the solid and the dashed curves.

Some of the structure of the curves in Figure 3.2a can be understood easily from theoretical models. If the demand is considerably less than the service potential, the expected queue length will be short, and the distribution of queue lengths will be approximately exponential. Under these circumstances, the standard deviation approximately equals the expected number in the queue. Thus, between time 0400 and 0900 and 2400 and 2800 (our 2800 can be interpreted as 0400 the next day) the curves in Figures 3.1 and 3.2a are approximately equal.

If for a prolonged period of time, the demand rate approximates the service capacity, the distribution of queue lengths should be uniform with a mean value of approximately one-half of the maximum queue length and a standard deviation of approximately one-quarter of the queue length. We can see from Figure 2.1 that the demand level reaches 55 operations per hour at approximately 0900 and remains between 50 and 60 till approximately 1400. From the period 1000 to 1400, the expected number in the queue lies between 10 and 15, which is

approximately one-half of 25, and the standard deviation of the estimate lies between 6 and 8, which is close to 25 divided by 4. The demand curve crosses the value of 55 operations per hour again between 2200 and 2300, but does not stay there long enough for the average or standard deviation even to approximate these equilibrium values. There is, however, a peak in the curve of standard deviation quite analogous to the shape of the curve between 0800 and 0900.

Also when the demand is considerably in excess of the service potential, the queue will be nearly saturated, with an exponential distribution in a reverse sense; that is, the probable number of places in the queue available to be filled is exponentially distributed. Under these circumstances, the standard deviation of the number in the queue should be approximately 25 minus the number in the queue. It is clear that this situation prevails between hours 1600 and 2200.

On this figure, the dotted lines represent the trivial situation of completely deterministic flow. These set one limit to the kind of service improvement which might be achieved by ideal metering and flow control. This question is discussed in more detail in Reference 4.

In fact, Reference 4 contains a number of other examples involving one other diurnal demand function and two other service rates. All of the inferences mentioned above are substantiated by these five other cases. Although the case of the single server queue is highly idealized, similar considerations prevail with multiple queues.

3.5 CONCLUSIONS

By a combination of numerical experiments and analysis, we have been led to the conclusion that fluctuations in the ratio of demand rate to potential service rate are responsible for the major time fluctuations in the parameters of queues of this type, including expected queue lengths, expected waiting time, total waiting time, standard deviations of the estimates, and other queue parameters. To the contrary, these queue parameters are relatively insensitive to details of server statistics. Therefore, in modeling situations with widely fluctuating demand and service rates, we can be rather free with our choice of service statistic models.

These results point to a secondary conclusion which must be qualified by further analysis: when the demand rate exceeds the service potential for any appreciable period of time, even complete regularization of flow does not eliminate delay. In general, any kind of congestion system experiences blockage when the demand exceeds the service potential, but the operating characteristics for demand rates somewhat below the service potential may be very strongly influenced by the protocol and statistics of the queues. In general, we can expect greater benefit from delay-reducing measures which increase the service rate than we can from those which change the shape of the response characteristic without significantly increasing the saturation rate.

4. OPERATIONAL MEASUREMENTS IN THE TERMINAL CONTROL AREA

4.1 STATEMENT OF PROBLEM

One of the primary problems facing analysts studying the air control system is lack of information regarding many aspects of the detailed operations of the system. For example, there is a little data available on the actual separation distances between incoming aircraft in the terminal area under IFR conditions. Therefore, since we do not know the statistical nature of this distribution, there are many important calculations that cannot be made. Simulations, such as those done at NAFEC, can sometimes be used to obtain such information, but very often there are psychological factors in the real situation which cannot be simulated. Consequently, whenever it is possible to do so, data should be acquired from operational situations.

One source of such data are the experiments which the FAA conducts from time to time. These involve the ATC system under operating conditions and are generally concerned with evaluating new equipment or new procedures. Tests of this type are, of course, very expensive, both in effort and in money, and it is important that they be designed to be as efficient as possible. This means that there should be appropriate hypotheses and a suitable experimental design. However, to the extent the essential objectives are not compromised, these tests could be a source of data which would be useful for purposes other than that of the test itself. In this section, we will identify a number of quantities relating to capacity in the terminal control area, the values of which do not appear to be very well established, but which are important to analysis, and therefore should be taken into account in any future operational testing to be done by the FAA.

The discussion which follows is intended to provide general principles for testing equipment and its modifications under operating conditions, and also to give some specific recommendations of what should be measured, why, and what information could be deduced from the results. These specific recommendations apply to the terminal control area of an airport operating under IFR conditions.

Before going into these matters, one general remark: The terms *capacity*, *delay*, *demand*, *safety*, *efficiency*, and others – have been sufficiently discussed in a previous report⁴ in their various practical contexts and in all details. In particular, a useful meaning of “capacity” was shown to be a response curve to a demand curve – a “functional transformation” (rather than a single number) analogous to the “transfer function” of electric networks, but non-linear. To improve air transportation, we submit that an understanding of *cause-and-effect relations* rather than further definitions of terms is the most profitable object.

4.2 GENERAL PROBLEMS IN OPERATIONAL TESTING

Turning to the general matter of drawing conclusions from trials, we further submit that – *if seen without guiding principles* – the situation may easily appear to be an impenetrable thicket of complexities. Even when analyzed by straightforward statistical methods, the situation is scarcely improved. Suppose, for example, that the various conceivable factors are listed, and the results of the actual trials in each combination have been found and tabulated. If there are f different factors describing the possible situations, an examination has to be made when each one has *at least* two different values – i.e., at least 2^f combined cases in all. Since each combination has to be tested a sufficient number of times (say s) to yield a statistically significant conclusion, we are faced with the necessity of at least $s \times 2^f$ trials. It is excessively conservative to take values of the order of $f = s = 10$. This gives a requirement of 10×2^{10} or 10,240 trials under controlled conditions. If each trial lasts one day, we have 10,240/365 or about 28 years of tests. Clearly, such a requirement is well beyond any practical usefulness.

What we have said simply repeats what scientists have known all along – that the ostensibly “hard-headed practical method of seeing how things really work out” can easily slip into absurdity. To be effective, operational tests have to be guided by *a priori* conceptions of the structure of what is being examined. These are obviously subject to alteration during the course of the tests, but they must always be present.

To repeat this point in another form: In the laboratory, it is possible to control most of the conditions of the matter under observation, so that one thing is examined at a time. Then the existing – or hypothesized – knowledge of how the different factors may interrelate can be used to guide any prediction of behavior of the complex system. But in *operational* tests, the blind collection of data can tell nothing: it must fit into a theoretical framework. A model is the current but less adequate rendering of this concept.

Finally, there is a last difficulty in attempting to draw conclusions of practical validity from masses of data. For example, it is impossible to record *everything* in any operation; one must merely select appropriate data. But without a preconception of the structural relationships of the system considered, it is quite impossible to tell what is appropriate. This fact, which has been constantly illustrated in operational testing in all the military and civilian applications of operations research since early in World War II – and even before – is but a comment on the practical limitation of a purely pragmatic approach: the accumulation of a vast “data bank.”

4.3 QUANTITIES TO BE MEASURED PRECISELY

It is useful to divide the things to be measured and recorded in a terminal control area into two classes: (1) those quantities which can be measured with needed accuracy, and (2) those which cannot and which therefore, if needed, will have to be calculated in terms of the former. In this section the quantities which can be – and should be – measured precisely are listed and the usefulness in obtaining them explained. The other factors that should be observed and recorded, of a more qualitative nature, will be given in the following section.

Quantities capable of precise measurement are the *times* at which certain events occur during the operation of an aircraft vis-à-vis the terminal. The following should be measured and recorded in a retrievable form:

- For Arrivals

T_A , the time at which the aircraft leaves the terminal entry fix on its way inward. This will be the same as its time of arriving at this fix if it is not held there; if this is the case, the time T_{AA} at which it reached this fix should be recorded. Then $T_A - T_{AA}$ is its time held.

T_B , the time of crossing outer marker
 T_C , the time of touchdown (down time)
 T_G , the time of reaching the gate

- For Departures

t_G , the time of leaving the gate
 t_C , the time of wheels up (off time)
 t_B , the time of release to en route controls

4.4 QUALITATIVE DATA TO BE RECORDED

For reasons to be given later, the numbers requested above can be used to derive important facts of the operation, from which the effects of various types and degrees of air traffic control can be studied. In addition to such data, it is also necessary, in varying degrees, to record the following data of a more qualitative sort accompanying each operation:

- Aircraft type and identity,
- Identification of terminal entry fix (for arrivals),

- Assigned runway number,
- Mode of approach of arrivals (ILS, NDB, etc.),
- Missed approach, if any (arrivals),
- Control mode used,
- General weather conditions, including degrees of visibility, wind velocity, and other relevant meteorological factors,
- State of traffic congestion at terminal, and
- Any unusual conditions that could make the operation untypical.

4.5 DEDUCTIONS

The first and most obvious deduction from the numbers gathered in Section 4.3 regards the "service rates" of landing and departing aircraft. Suppose that during certain periods of the day, there is a high traffic density, so that the number of aircraft landing (down times per hour) represents the rate at which landings occur when the landing facilities are fully used. Then the number of landings per hour, as well as the distribution of time intervals between one landing and the next, gives an important parameter from which much information bearing on capacity can be derived by use of existing analytical tools. (It is the quantity μ in the queuing and related phenomena set forth in a companion report.³) Similar remarks apply to the case of departing aircraft.

5. STRUCTURING COST-BENEFIT METHODOLOGY FOR ATC CAPACITY IMPROVEMENT MEASURES

5.1 STATEMENT OF THE PROBLEM AND DESCRIPTION OF RESULTS

One of the primary objectives of the current study is to provide the FAA with the tools necessary to perform a cost/benefit analysis on the measures proposed to increase ATC system capacity. Of course, there is no simple universal tool which can be applied mechanically throughout the entire spectrum of potential improvements. Rather, the evaluation of each kind of improvement will require its own and specific analytical basis. For example, in the case of adding more airports to the system, the equations would differ from those used for improved precision in final approach spacing. However, this does not necessarily mean that every different kind of improvement requires a totally different approach for its evaluation. In fact, from an analytical point of view, most of the measures can be organized into a relatively small number of classes, each class consisting of those measures which would be evaluated from a common point of view. Thus, the number of different approaches required is far less than would be suggested by the number of potential measures to be considered. The purpose of this section is to perform this classification.

The work in this chapter is based on a companion report¹ which contains an extensive tabulation of ATC factors related to capacity and efficiency. Our procedure was to consider each factor separately and identify the system characteristics affected (both positive and negative), the nature of the improvement, and the key effects which the measure would have on system bottlenecks. The majority of measures involved the terminal area; consequently, the principal bottlenecks were runway clearance time, effective approach separation interval, and the controller unit capacity. The results are presented both graphically and in tabular form. In addition, an example is provided of the steps involved in making the cost/benefit determinations for such capacity improvement measures.

5.2 HISTORICAL BACKGROUND

In the early phases of this contract, we were required to establish certain terms and measures, particularly to give precise meaning to the concepts of capacity, demand, and delay. These were intended to serve us in appraising and stating in what respects and to what extent the air transportation system is performing its function, now and in future projections; and to guide in setting up precise objectives for its improvement. Such definitions have been given^{1, 4, 5} and were accompanied by an extensive supporting discussion. As a background, the general procedural principle was given, excerpted as follows from Chapter 3 of Reference 4:

"The process of defining terms and measures is iterative. Relevance, intelligibility, and measurability are the criteria for the choice of a first tentative definition. To improve on these definitions, it is necessary to form a precise conception of the mechanism underlying the system under consideration, e.g., the ATC component, the air terminal, and the like. This step is often called 'setting up a model.' After enough examinations of this sort, the degree of adequacy of these terms and measures to express organic features of the situation become better understood. Moreover, certain other factors may have been missed in the initial formulation of terms and measures. With the improved list, more relevant models can be set up and operated. This interplay of terms and measures with models and measurements – this process of cyclical refinement – is common to all developments of science and its technological applications."

While in a general way the term "capacity" is used to express the capability of the system to supply a required level of service, a more detailed analysis has shown that it represents a response curve (of service rendered) to a demand curve. It is comparable to the "transfer function" in electric circuitry, but is strongly non-linear. It is also affected by the statistical fluctuations of the situation, and hence requires methods of probability for its calculation.

In later work on this contract, we presented methods for calculating capacity, as well as delay and similar deteriorations of service, under widely varying conditions.³ In a more recent report,¹ we sought to examine all of the general and many of the specific actions or technological advances that might lead to improvements in the ability of the ATC system to meet demands.

In this chapter, we attempt to structure the material of that report to make it more useful as a means for studying the capacity question. The premises under which we worked represent a continuation of the process described earlier in this section and include:

- The reason why a measure of ATC system capacity is sought is to provide a tool that can be used to evaluate the effectiveness and cost-effectiveness of various measures which might be undertaken by the FAA, or by the airlines, in efforts to increase this capacity over current levels.
- Capacity measurements only become meaningful and useful when related to demand; accordingly, the analysis of capacity changes is significant only in situations in which demand is, or will be, pressing against current limitations of capability.

- For all practical purposes, the changes which might be made to the present system are marginal; in all areas where demand is high and current capacity is, or will be, stretched, the opportunity to effect fundamental changes is practically non-existent. Near-term progress can only be made by a series of relatively small additions, improvements, changes, or adjustments.
- The ATC system, relative to a given local demand and growth in demand, is composed of a current or "next" bottleneck, plus a series of other potential bottlenecks, one of which will take over the limiting function whenever the current or next bottleneck is sufficiently improved.

The picture which flows from the above premises is that our work should seek not a single capacity measure, but the development of a tool for testing the impact of specific possible changes as they apply to existing and potential bottlenecks. There is general acceptance, for example, that *airways* capacity does not, and will not in the near future, impose a bottleneck and that, in the limited areas where difficulties may now be encountered (transatlantic airways, airways intersections, etc.), rather simple techniques could be employed to achieve relief. This area, therefore, merits little current analytical attention. In fact, every study^{6,7} – including our own – rapidly focuses on capacity near airports and during final approach and landing.

Over a period of time, we have gradually concluded that there are three limiting and fundamental characteristics which, in a given local situation, may impose the local capacity bottleneck. They are:

- Runway clearance time;
- Effective approach separation interval; and
- Controller unit capacity (taking into account interfaces among control activities).

The first two of these characteristics are based on limits that may be reduced by technological factors but which, at a given time, cannot be exceeded without at least some loss of safety. The third item differs in that control units can be added "as necessary," but only at high and increasing cost per unit of effective work performed.

5.3 SYSTEM IMPROVEMENTS AND THEIR EFFECTIVENESS

In the following discussion we will describe the steps required to identify the path by which we can, starting with a potential technological or procedural

change, (1) determine its effect upon capacity as an intermediate point, and (2) ultimately find the worth of the capacity change to those groups of ATC users and the public which would be affected by the change.

A more complete picture of the process by which system improvements can be evaluated involves other steps between these two points. A simple example is given in Figure 5.1.

- *The Starting Point.* As implied above, this is a technological innovation or procedural change. In the example of Figure 5.1 the technological advance shown is the employment of a microwave multibeam ILS.
- *The System Characteristics Influenced.* The system characteristics influenced by the improved ILS is the final approach path guidance. Note that any one of a number of other technological improvements or procedural steps might influence this basic system characteristic, each one measurable in terms of the degree of improvement possible and the costs involved in obtaining this improvement.
- *Intermediate Improvement Mechanism.* The improved ILS will permit controllers to employ variable glide angles and thereby shorten the length of approach to the airport runways. (In some cases this intermediate step is unnecessary; the system characteristic will impact directly on the bottleneck factor.)
- *Bottleneck Factor Influenced.* The employment of shorter approaches adaptable to the natural glide characteristics of aircraft can permit a reduction in the separation minima employed – and thereby increase the landing rate. In the process, however, the controller will have a more complicated function to perform, so there may be a side effect of increasing controller workload per aircraft landed. Thus, in this case, two bottleneck parameters are influenced, one favorably and one unfavorably.
- *Operational Results.* Having determined the effect of the change in capacity which results from the modification to the bottleneck factor established, the next step is to determine the operational result of this change when applied to a specific airport. In the specific improvement illustrated in Figure 5.1, this operational result can be expressed either in increased airport throughput or in decreased delays.

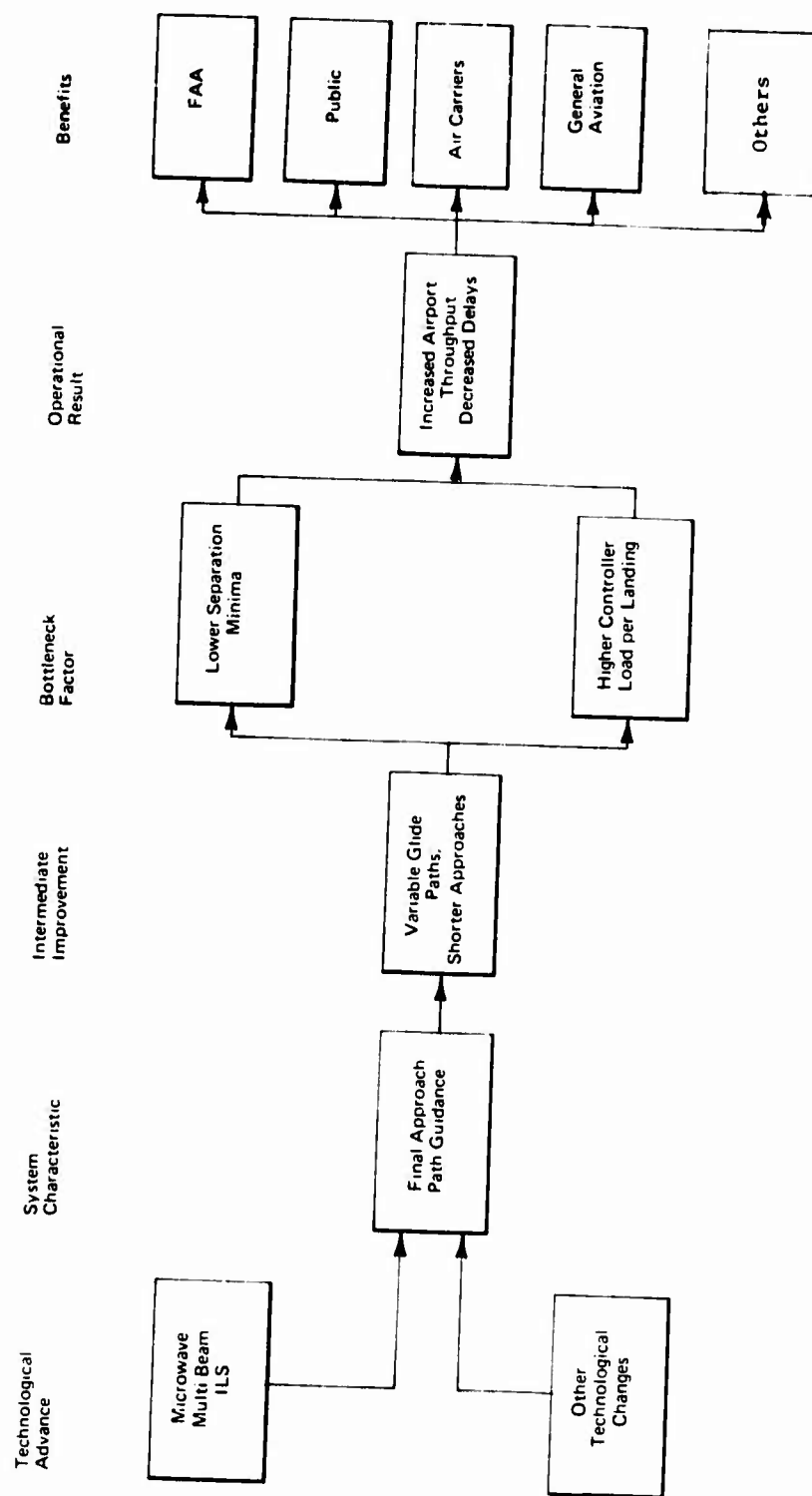


FIGURE 5.1 AIR TRAFFIC CONTROL IMPROVEMENT PATTERN

- *Benefits.* In the final analysis, we will have to evaluate improvements in terms of the benefits to the various users of the system. There are two reasons why this must be done: (1) this is the only way we can determine the worth of the various technological advances to the system from a cost/benefit point of view; and (2) this mechanism allows us to compare the worth of competing improvements on a relative basis.

In general, we may regard that part of the block diagram between "system characteristic" and "bottleneck factor" as a transfer function, with a structure that will apply to almost any airport. Therefore, our first task might be to determine what these transfer functions are for the various system characteristics. In fact, we have already established a number of them. The next step is an engineering one, namely, to identify the characteristics of the technological advances in terms of the parameters which affect the system characteristics. The third job is an operational and economic one which will vary from airport to airport and consists of three parts: (1) to determine the investment and annual operating costs implied by the technological advance when applied at that particular airfield, (2) to evaluate the operational consequences of the technological advance in terms of the specific improvement of the operational variables involved which reflect capacity, and (3) to estimate (in dollar terms, if possible) the consequences of these changes on the users of the system.

In the remainder of this chapter, we will describe our attempt to illustrate by diagrams a formal structure of the means whereby potential improvements to the ATC system can be analyzed. While the specific measurement of capacity is avoided, the capacity concept is implicit in the process. The "system characteristic" changes and their relationships to the critical "bottleneck factors" are indicated, along with the interaction path between the two. The operational result and benefit steps are omitted because of redundancy in the case of the benefits, and also because they can usually be identified from the nature of the specific bottleneck factor involved. In summary, our general approach is that of breaking down the means whereby improvements might be made into a series of steps which will permit:

- Measurement of the improvement that a system characteristic can derive from a technological change or procedural improvement; and
- Establishment of a pattern (perhaps quite generalizable) which relates changes in system characteristics with changes in bottleneck characteristics.

In other words, we are attempting to set up a limited, but all-encompassing, structure which, with further development, may permit relatively direct evaluation of the impact of possible improvements.

5.4 RESULTS

Figures 5.2, 5.3, and 5.4 present generalized block diagrams showing the interactions between system characteristics and bottleneck parameters. In creating these charts, a large number of the items mentioned in Reference 1 was eliminated (as listed in Table 5.1), since they were considered fundamentally redundant (usually representing slightly different perspectives or processes which were fundamentally similar). Another group of items from Reference 1 (Table 5.2) was dropped from consideration as not being structurally relevant to the capacity question — although they did deal with subjects influencing ATC system costs, or non-ATC system-related mechanisms for relief of capacity problems. The residual items were then grouped in terms of which of the various bottleneck factors they influenced, either singly or simultaneously. To simplify the presentation, three diagrams were prepared relating to airports, aircraft, and the ATC system, respectively.

Finally, Table 5.3 shows the intermediate step (the interaction path) not shown in Figures 5.2, 5.3, and 5.4 for each of the improvements. Table 5.1 lists the items we felt were reasonably equivalent to other items and, therefore, did not have to be shown specifically. On the table the particular item of this category is listed on the left and the equivalent on the right.

Three items shown in Table 5.3 are not shown in Figures 5.2, 5.3, and 5.4. They all contribute to more efficient use of resources and are as follows: 2243, Diversified Communication Link; 2411, Spread Peak Load; and 2461, Improve Load Factor.

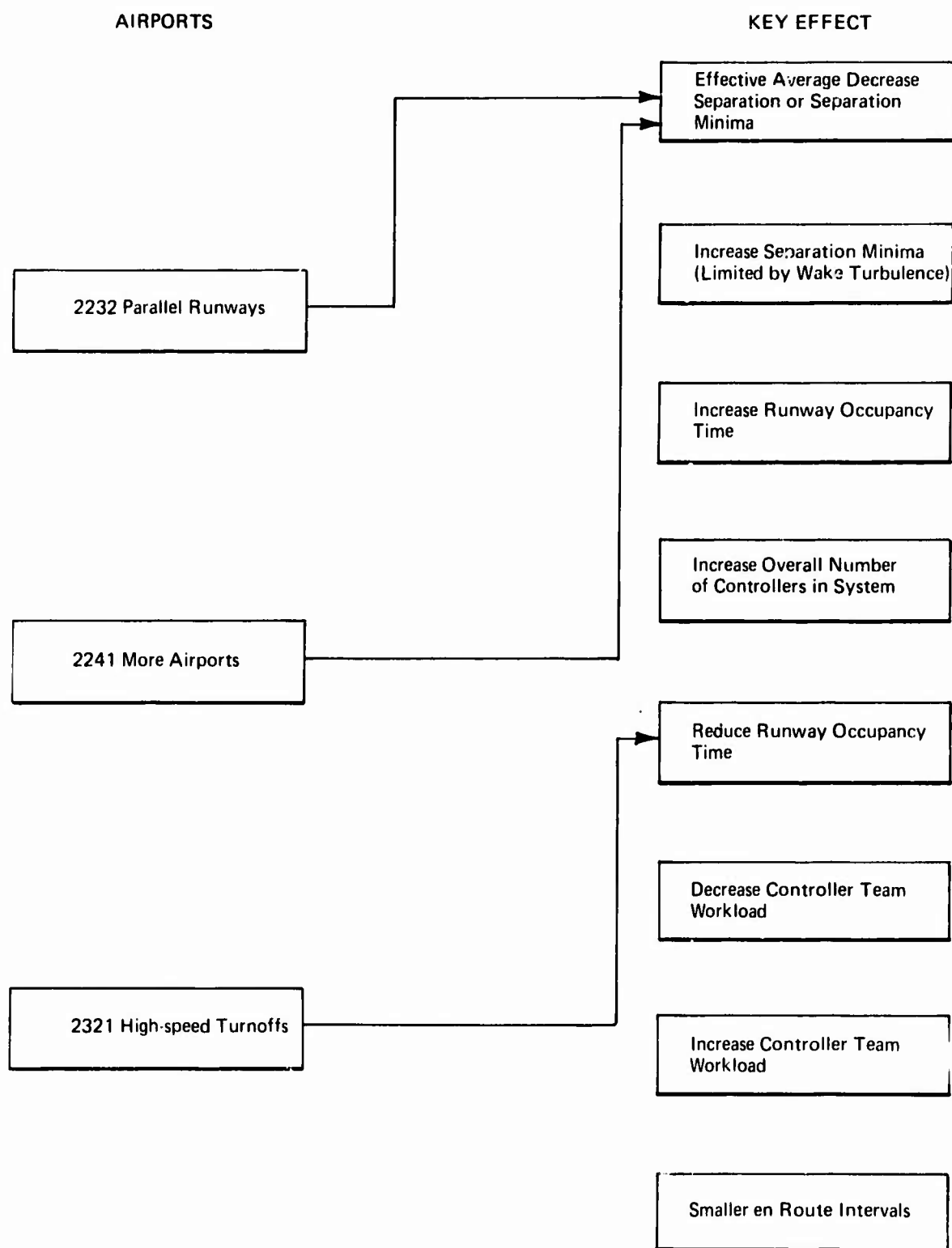


FIGURE 5.2 LIMITING AIRPORT CHARACTERISTICS RELATED TO KEY EFFECTS

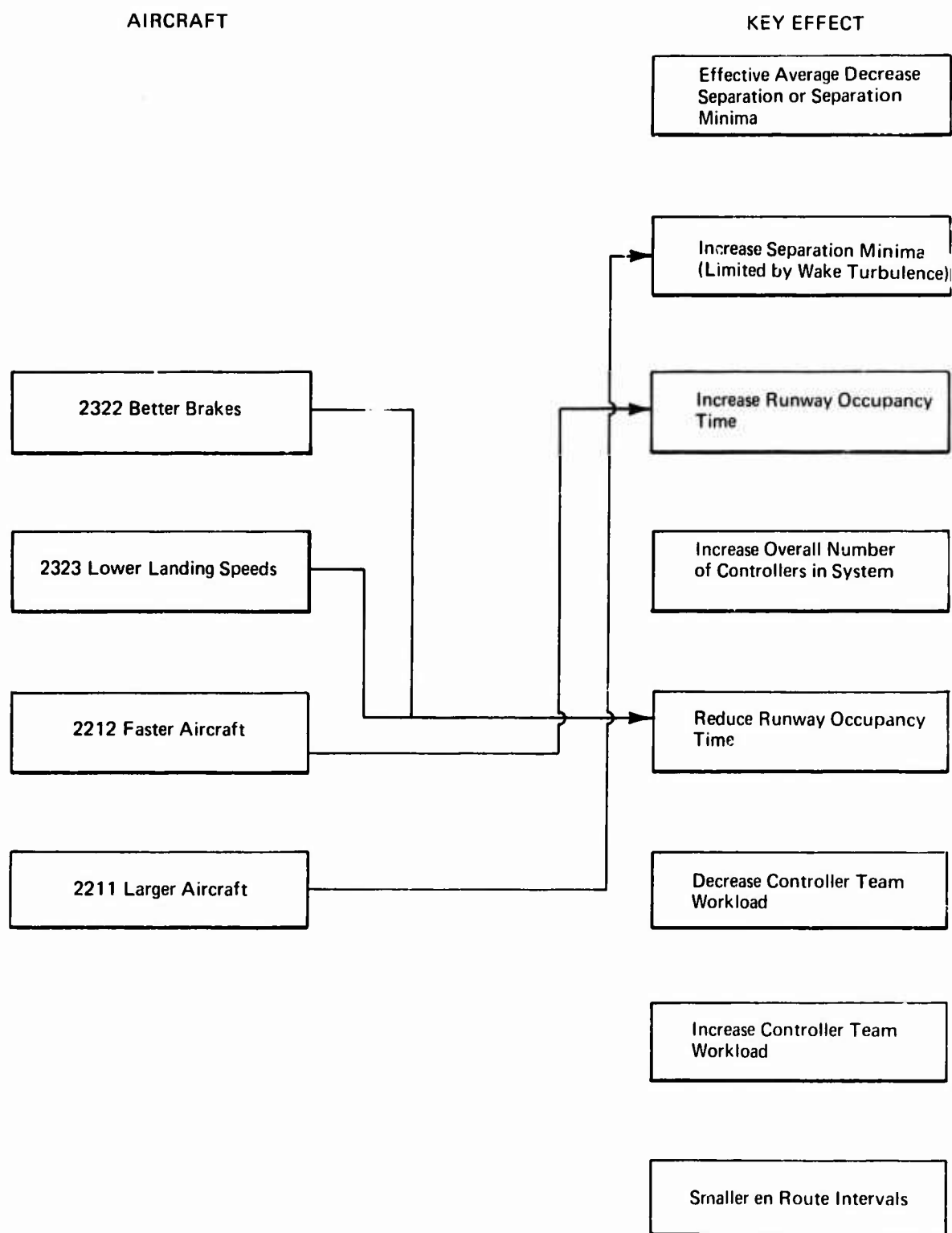


FIGURE 5.3 LIMITING AIRCRAFT CHARACTERISTICS RELATED TO KEY EFFECTS

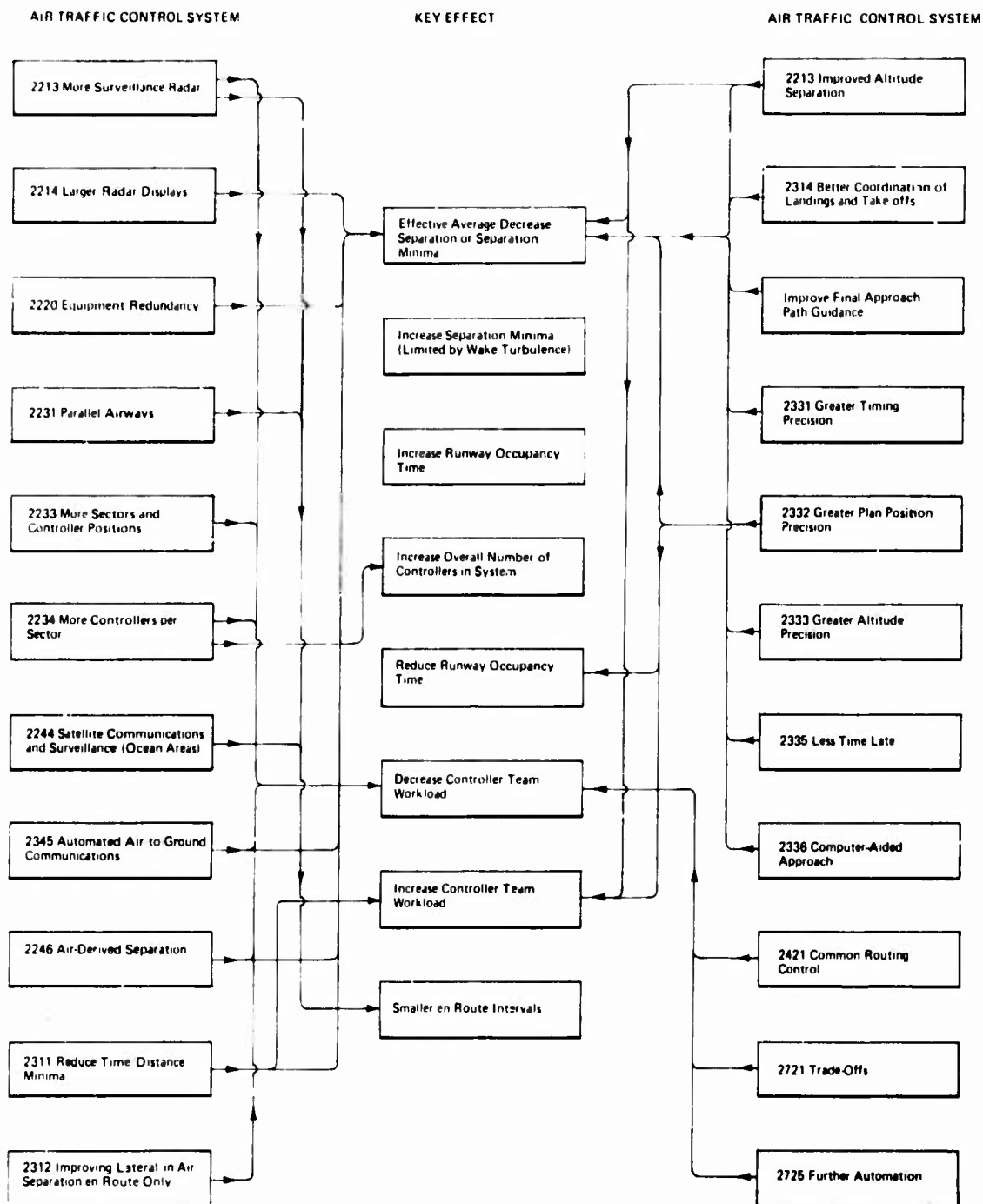


FIGURE 5.4 LIMITING ATC CHARACTERISTICS RELATED TO KEY EFFECTS

TABLE 5.1

CAPACITY-INFLUENCING ITEMS CONSIDERED AS EQUIVALENT TO OTHER ITEMS

Items	Equivalent Items
More Surveillance Radar Displays	Equipment Redundancy
More Communication Channels	Equipment Redundancy
Dual Runways	Parallel Runways
Special Requirements for High Precision	Length of Final Approach Path
Improved Radar Displays	<div> <div>Reduce Time-Distance Minima</div> <div>Lateral in-Air Separation</div> <div>Time Late</div> </div>
Metering	<div>Computer Aided Approach</div> <div>Sequencing and Metering</div>
Sequencing	<div>Computer Aided Approach</div> <div>Sequencing and Metering</div>
Regularization of Flow	<div>Computer Aided Approach</div> <div>Sequencing and Metering</div>

TABLE 5.2

CAPACITY-INFLUENCING ITEMS OMITTED FOR CURRENT STRUCTURAL IRRELEVANCY

VTOL and STOL Aircraft
 Common Reference Standards
 Intersections and Other High-Density Traffic Points
 Interface Management
 Remote Data Transmission and Display
 Communication Channel Allocations
 Higher Reliability
 Reserve Responses to Rare Excesses
 Weather
 Reserve Responses to Equipment Failure
 Planning for Graceful Degradation
 Reserve Response
 Trade-Offs

TABLE 5.3

**RELATIONSHIPS BETWEEN ATC IMPROVEMENT MEASURES AND
THEIR KEY EFFECTS UPON THE SYSTEM**

Improvement Measure	Interaction Path	Key Effect
Larger Aircraft	Wake Disturbance	Increase Separation
Faster Aircraft	Higher Landing Speed	Increase Runway Occupancy Times
More Surveillance Radar Coverage	Reduce Procedural Control in some Areas	Smaller en Route Interval Lower Controller Workloads
Larger Radar Displays	Better Measurement of A/C Separation	Decrease Separation Minima
Equipment Redundancy	Lower Risks due to Equipment Failure	Decrease Separation Minima
Parallel Airways	Decrease Traffic at Intersections	Increase en Route Capacity
Parallel Runways	More Efficient Use of Terminal Area Airspace	Decrease Separation Interval
More Sectors and Con- troller Positions	Fewer Operations per Controller Per Unit Time	Ease Controller Workload
More Controllers per Sector	Fewer Operations per Controller per Unit Time	Ease Controller Workload per Sector Increase Number of Overall Con- trollers in System
More Airports	More Approach Paths and Runways	Reduce Separation Minima
Diversified Communication Links	Better Utilization of Technology	Reduce Systems Costs/Unit Capacity
Satellite Communications and Surveillance	Extends Localization Coverage	Decrease en Route Separation Minima in Ocean Area
Automated Air-to-Ground Communications	Automation of Data Handling	Lower Controller Workload (per operation) Decrease Separation Minima
Air-Derived Separation	More Efficient Use of Terminal Area Airspace	Lower Controller Workload (per operation) Decrease Separation Minima
Reduce Time or Distance Minima Between Aircraft on the Same Track	Spacing Reduction (limited by wake)	Increase Controller Workload Decrease Separation Minima

TABLE 5.3 (Cont.)

Improvement Measure	Interaction Path	Key Effect
Improving Lateral in-Air Separation	Simplify Conflict Resolution	Reduce Controller Workload (en route only)
Improved Altitude Separation	Increase Numbers of Non-Conflicting Flight Levels	Decrease Separation Minima Increase Controller Workload
Better Coordinated Landings and Take-Offs	More Efficient Use of Runways and the Terminal Area Airspace	Decrease Separation Minima
Length of Final Approach Path	Shorter Glide Slopes	Reduce Runway Occupancy Time Reduce Separation Minima
High-Speed Turnoffs	Leave Runway at Higher Speed	Lower Runway Occupancy Time
Better Brakes	More Rapid Deceleration on Runway	Lower Runway Occupancy Time
Lower Landing Speed		Lower Runway Occupancy Time
Greater Timing Precision	More Efficient Use of Time	Lower Separation Minima
Greater Plan Position Precision	Easier Use of Multiple Runways	Lower Separation Minima Lower Runway Occupancy Time Higher Controller Workload per Operation
Greater Altitude Precision	More Use of Twisted Paths	Lower Separation Minima
Less Time Late	More Efficient Use of Time in Final Approach	Less Separation Minima
Computer-Aided Approach Sequencing and Metering	Reduce Separation to Minimal Sequencing to Minimize Average Spacing	Lower Separation Minima
Spread of Peak Hour Load	Reduce Congestion Peaks	Improve Resource Allocation
Replace Step-by-Step Routing with Common Control	Better Arrival Scheduling	Lower Controller Workload
Mergers, Restrictions on Route Competition	Increased Load Factor and Reduced Numbers of Operations	Fewer Operations at Peak Hours
Trade-Offs with Lower Workload for Controller or Pilot	More Control Activities by Pilot	Reduce Controller Workload

6. ESTABLISHING PRACTICAL SAFETY GOALS IN AIR TRAFFIC CONTROL

6.1 STATEMENT OF THE PROBLEM

Air traffic density and safety are inversely related under many circumstances. Graham and Orr⁸ have shown that collision exposure in terminal areas is strongly related to a traffic factor which is a weighted measure of the amount of traffic. The concept of preserving the safety of aircraft by spacing them far apart has been formalized in separation standards, operating rules, and the structure of airways and terminal air space. Thus, many capacity-limiting constraints have been built into the fabric of air traffic control for the purpose of fostering safety.

The relationship between minimum spacing or layout of paths and the maximum possible number of operations is always determinable in principle, and many cases have been worked out.⁹ In most cases, however, the relation to safety has not been worked out quantitatively. Therefore, safety does not enter into the specification of ATC system capacity explicitly, but enters implicitly through the various standards, rules, procedures, and designs which are explicit. This is not to say that safety is inadequate or that it is handled irrationally. On the contrary, there is ample evidence that safety has improved steadily for the last half century, and that the level of risk in U.S. commercial aviation is both very low and under good control when viewed as a system variable. On the other hand, it is difficult to make cost-benefit analyses or to make comparisons of system alternatives where one of the benefits is a reduction in risk where there is no opportunity to estimate how much this risk reduction might be. There seem to be two reasons for this difficulty: (1) the quantitative relationship between safety and operating parameters, such as minimum separation standards, is often extremely difficult to analyze (this, however, is the lesser cause); and (2) a more important cause, the confusion between striving for the ideal of no risk and accepting no risk as a criterion for satisfactory performance of the air transportation system. Adopting no risk as a criterion of satisfactory performance, even for reasons of idealism, not only guarantees failure to meet the criteria, but deprives us of the benefits of rational risk analysis. The benefits will arise from discoveries and insights about real differences in safety, from the results of trade-off studies which show how much the present level of safety costs us and how we might get more safety possibly for even less expenditure, and by making explicit the promises and consequences of decisions concerning safety which might otherwise be obscured.

There are many reasons why the adoption of a finite level risk as a criterion of minimum of satisfactory performance is difficult. In the first place, outside of military systems analysis, there is a great reluctance to discuss the role which human beings play in increasing or decreasing each other's risks. This reluctance can be compared with the reluctance of the previous generations to discuss cancer

Preceding page blank

or tuberculosis, which might have been an impediment to rational treatment and research. Also, no unit for the measurement of air transportation risk has been generally accepted. Finally, no quantitative thresholds for a minimum of acceptable level of safety have been proposed.

In a separate report² we have raised – and partly answered – the question: “How could practical finite ATC safety goals be established?” Inasmuch as this report is generally available, its contents will be reviewed here only briefly.

6.2 METHOD OF APPROACH

The key step in our solution of this problem is the adoption of fatality risk per unit of exposure time as a measure of safety. This unit has been shown^{1, 0} to be particularly useful in analyzing human social responses to risks from a variety of sources, and therefore to trades in which risk is one component. We have also shown in a previous report⁴ that many aviation accident statistics can be easily graphed and compared when described in terms of this unit.

The next step was to estimate the contribution to air traffic risk from a number of different causes. In the safest category of aviation – domestic scheduled air carrier flights – there is a serious conflict between validity of observational data and detail because fatal accidents are so infrequent. In a compromise, 36 classes of fatal accidents were developed from six phases of flight and six type-cause categories.

The overall level of risk in this category of aviation is easy to determine with fair precision. It is approximately 1.6 fatalities per million hours of exposure. This is comparable to risks which are accepted by large numbers of people in exchange for economic benefits or other satisfactions.^{1, 0} We did not investigate the questions of whether this level of safety is near a rational optimum or whether it will continue to be acceptable when, for example, air travel greatly increases. Instead, we examined the implications of choosing a substantially lower risk level, one fatality per 10,000,000 hours of exposure, as a design standard. Adoption of such a standard calls for a 16-fold reduction in the overall rate of fatalities, and its achievement would therefore represent a very great improvement over present-day performance.

Any attempt to achieve this level by improving air traffic control alone is doomed to failure. By the most generous reckoning, less than half of the fatal accidents in our sample could be associated in any way with air traffic control. On the other hand, the goal could be reached by a proportional reduction in risk from all causes. This produces a mutually compatible set of risk goals for each of the 36 classes. These could be used, in turn, as the basis for safety and reliability budgets suitable for subsystem and equipment design.

6.3 CONCLUSIONS

This work has led to four conclusions:

1. The content and quantity of available data are sufficient for this process to be carried out;
2. The incidence of various classes of accidents is consistent with the use of fatality per hour of exposure as a unit of measurement, and with the assumption of statistical independence among accidents over a moderate period of time;
3. Passenger and crew risk, in these units, can be reduced to levels characteristic of every human activity perceived as "negligibly risky," only if sources of risk totally outside of air traffic control are substantially reduced, as well as risks with a possible relation to air traffic control; and
4. The number of incidents implied by the resulting safety targets is so small that an indirect, rather than a direct, estimation of safety performance would be required to validate any attempts to achieve them.

REFERENCES

1. Raisbeck, G., Factors Influencing Capacity and Efficiency in Air Traffic Control, Report No. FAA-RD-71-107, Arthur D. Little, Inc., November 1971.
2. Lister, S.F., and Raisbeck, G., An Approach to the Establishment of Practical Air Traffic Control Safety Goals, Report No. FAA-RD-71-36, Arthur D. Little, Inc., May 1971.
3. Koopman, B.O., Analytical Tools for the Study of Airport Congestion, Report No. FAA-RD-71-55, Arthur D. Little, Inc., September 1971.
4. Raisbeck, G., et al., A Study of Air Traffic Control System Capacity, Report No. FAA-RD-70-70, Arthur D. Little, Inc., October 1970.
5. Raisbeck, G., Everett, J.L., and Koopman, B.O., Continuing Studies of Air Traffic Control System Capacity, 1970-1971, Report No. FAA-RD-72-2, Arthur D. Little, Inc., January 1972.
6. Report of the Department of Transportation Air Traffic Control Advisory Committee (2 Vols.), Supt. of Documents, U.S. Government Printing Office, December 1969.
7. Proc. IEEE, Special Issue on Air Traffic Control, Vol. 58, No. 3, March 1970.
8. Graham, W., and Orr, R.H., Terminal Air Traffic Flow and Collision Exposure, Proc. IEEE, Vol. 58, No. 3, March 1970, pp. 328-336.
9. Department of Transportation Air Traffic Control Advisory Committee Report, Vol. 1, Sec. 3.2.2, op. cit.
10. Social Benefit vs. Technological Risk, Chauncey Starr, Science, Vol. 165, No. 3899, September 19, 1969, pp. 1323-1238.

Preceding page blank